# Vectorisation and tf-idf analysis